REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
18-11-2016	Final Report	8-Dec-2014 - 7-Jul-2016

4. TITLE AND SUBTITLE
Final Report: A Heterogeneous High-Performance System for Computational and Computer Science

5a. CONTRACT NUMBER
W911NF-15-1-0023

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
106012

6. AUTHORS
Hoda El-Sayed, George Ude, Bo Yang

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Bowie State University
14000 Jericho Park Road

Bowie, MD 20715 -9465

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
66387-CS-REP.1

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This DoD HBC/MI Equipment/Instrumentation grant was awarded in October 2014 for the purchase of a supercomputer. The equipment that considered was a Cray XC40 supercomputer to restore Bowie State University to provide the BSU students with access to state of the art High-Performance Computing (HPC) tools that can be integrated with existing curricula and support our research to modernize and dramatically advance our research and educational programs. The supercomputer will support our institution in expanding our interdisciplinary research and education across many departments for the benefit of our students and our faculty. It will help in enhancing the

15. SUBJECT TERMS
Cray XC40, HPC research

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Hoda El-Sayed
UU	UU	UU	UU		19b. TELEPHONE NUMBER 301-860-3964

Report Title

Final Report: A Heterogeneous High-Performance System for Computational and Computer Science

ABSTRACT

This DoD HBC/MI Equipment/Instrumentation grant was awarded in October 2014 for the purchase of a supercomputer. The equipment that considered was a Cray XC40 supercomputer to restore Bowie State University to provide the BSU students with access to state of the art High-Performance Computing (HPC) tools that can be integrated with existing curricula and support our research to modernize and dramatically advance our research and educational programs. The supercomputer will support our institution in expanding our interdisciplinary research and education across many departments for the benefit or our students and our faculty. It will help in enhancing the High Performance Computing (HPC) course taught in the department of computer science as to attract more graduate students from many disciplines where their research involves HPC.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received Paper

TOTAL:

Number of Manuscripts:

Books

Received Book

TOTAL:

Received	Book Chapter

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Post Doctorates

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Faculty Supported

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED

FTE Equivalent:
Total Number:

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Technology Transfer

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(15-11-2016)*	**2. REPORT TYPE** **Final Report**	**3. DATES COVERED** *(From - To)* 12/8/2014 – 9/29/2017

4. TITLE AND SUBTITLE

A Heterogeneous High-Performance System for Computational and Computer Science

5a. CONTRACT NUMBER
W911NF-15-1-0023

5b. GRANT NUMBER
66387-RT-REP

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Hoda El-Sayed, PI and Dr. George Ude, Co-PI

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Bowie State University
14000 Jericho Park Rd
Bowie, MD 20715-9465

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

ARMY RESEARCH OFFICE
ARO-RTP, 4300 S. MIAMI BLVD
DURHAM, NC 27703

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This DoD HBC/MI Equipment/Instrumentation grant was awarded in October 2014 for the purchase of a supercomputer. The equipment that considered was a Cray XC40 supercomputer to restore Bowie State University to provide the BSU students with access to state of the art High-Performance Computing (HPC) tools that can be integrated with existing curricula and support our research to modernize and dramatically advance our research and educational programs. The supercomputer will support our institution in expanding our interdisciplinary research and education across many departments for the benefit or our students and our faculty. It will help in enhancing the High Performance Computing (HPC) course taught in the department of computer science as to attract more graduate students from many disciplines where their research involves HPC.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Hoda El-Sayed

a. REPORT	b. ABSTRACT	c. THIS PAGE		5	19b. TELEPHONE NUMBER *(include area code)* (301)860-3964 REPORT DOCUMENTATION PAGE (Standard Form 298)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Proposal No: 66387-RT-REP

A Heterogeneous High-Performance System for Computational and Computer Science

Summary

The goal of the grant awarded is to acquire a high-performance computing instrument to support an interdisciplinary team of research faculty from the departments of computer science and natural science at Bowie State University. The supercomputer is not only to expand the research infrastructure at the institution but also to enhance the high-performance computing training provided to both undergraduate and graduate students. The Cray XC40 is ideal for our research in the Department of Computer Science, where research is considering parallel programming models productivity, and in particular the promise and problems of the Partitioned Global Address Space (PGAS) model as well the productivity of GPU accelerated HPC systems. The supercomputer is also ideal for the research conducted in the Department of Natural Science, as research faculty work on research to sequence large databases of DNA through a DNA Barcoding Initiative to sample, identify and classify species. In addition to research, the supercomputer will also be used to enhance the educational experience of our students in many classes and programs. The Cray XC40 will allow assigning realistic computational problems that can integrate research and teaching in the STEM disciplines. Next Spring semester, The Principle Investigator offers a High-Performance Computing graduate course. The supercomputer will also be used to enhance the educational experience of our students enrolled in the course.

A workshop was conducted last month to train and facilitate the usage of the supercomputer to the faculty and students. Other workshops will be conducted to show faculty and students how to stimulate research in engineering, science and mathematics through computational modeling and simulations. Work will be widely disseminated through standard academic venues.

This DoD Research and Education Program for Historically Black Colleges and Universities and Minority-Serving Institutions (HBC/MI) Equipment/Instrumentation grant was awarded in October 2014 for the purchase of a supercomputer. The equipment awarded was a Cray XC40 supercomputer to restore Bowie State University to provide the BSU students with access to state of the art High-Performance Computing (HPC) tools that can be integrated with existing curricula and support our research to modernize and dramatically advance our research and educational programs. It took around one year of considering different options to get the best system including negotiation and installation of the Cray system as it was customized to fit our research needs in the university. Finally, the supercomputer was installed in December 2015. The networking configuration and many site preparations and solving many unforeseen issues took around six months as the University had to conduct a search and appoint a new system administrator. Finally, in July 2016, the networking was configured and the system was ready for use. The system, which was named Sphinx to symbolize speed and intelligence, has a total of 12740 processing cores capable of performing at 59 TeraFLOPS (or 59 trillion calculations per second). Sphinx is a heterogeneous system with a blend of the most advanced processing technologies including Intel Haswell multicore chips, Intel Phi manycore chips and NVIDIA GPUs to provide our researchers with a range of options. CHIP – Center for High-Performance Information Processing is the name our newly established supercomputer center.

Sphinx will support our institution in expanding our interdisciplinary research and education across many departments for the benefit or our students and our faculty. It will help in enhancing the High Performance Computing (HPC) course taught in the department of computer science as to attract more graduate students from many disciplines where their research involves HPC. It will also help the undergraduate students be more aware of HPC concepts such as HPC simulations and data analytics, and to apply them as powerful tools in their work. As high performance computing and computational science became a critical research investigation tool in the fields of chemistry and natural sciences, it is essential that our students and faculty develop a great deal of knowledge of HPC. Various interdisciplinary research efforts are launched at Bowie State University, most of which requires the use of the proposed instrument to improve the research productivity. Sphinx is a high-performance computing system composed of an integrated GPU-based Parallel Computer and a Storage Area Network for short term data storage to support interdisciplinary research by faculty from the departments of computer science and natural science. This system will not only expand the research infrastructure at the institution but will also enable high-performance computing training for both undergraduate and graduate students. The Center for High-Performance Information Processing (CHIP), will support high-performance computational science research and education, with emphasis on computational biology, computational chemistry and parallel computing research. In addition, CHIP, will work on integrating parallel computing concepts into existing curricula as per the IEEE Technical Committee on Parallel Processing recommendations. There are different research studies being

developed at Bowie State University through multiple faculty in various disciplines in both Departments, Computer Science and Natural Science.

In order to facilitate the usage of the supercomputer, a one-day workshop was conducted by the Principal Investigator and Cray representatives in October 2016. The purpose of the workshop was to provide all the necessary information and basic hands-on experience needed for any faculty or student to be able to use the supercomputer. During the first session of the workshop, a brief explanation of high performance computing and the description of the system was presented. In addition, different examples where a supercomputer became a necessity for nowadays research were given. Faculty from different departments in the university as well as undergraduate and graduate students attended the workshop. In the second session of the workshop, the attendees were exposed to the usage of the supercomputer. They all were able to log on the system and learned how to compile on the Haswell, the Phi, and the NVIDIA GPUs. Small programming examples were given and the attendees were able to test them on the supercomputer. Everyone was impressed with the workshop and the students got so interested in conducting their research using it.

The Principle Investigator offers a High-Performance Computing graduate course. The supercomputer will also be used to enhance the educational experience of our students enrolled in the course. Students will be given accounts on the supercomputer and will have to use it to run all their parallel programs that they will be assigned during the course. In addition, several graduate students now have accounts on the new supercomputer and will start using it in their research.

Through Sphinx, we will carry out research in parallel programming models productivity, including the Partitioned Global Address Space (PGAS) model. We will also examine the productivity of GPU accelerated HPC systems. In addition, our research team will be able to sequence large databases of DNA through a DNA Barcoding Initiative to sample, identify and classify species. DNA barcoding is a new tool for identifying biological specimens and managing species diversity. It provides a way to identify and study medicinal plants that exist world-wide and have never been studied before. Another research area to be enabled by this HPC system is computer forensics. Explosion in data (big data) generating applications, advances in cloud computing, supercomputing, and the availability of cheap memory and storage led to enormous amounts of data to be sifted through in forensic analysis. This can be very critical for time sensitive investigations especially for governmental or industrial organizations. Using the supercomputer, we will develop new techniques for data organization and for providing the needed analysis in a timely fashion.

The Principal Investigator has conducted research work in heterogeneous computing using General Purpose Graphical Processing Units (GP GPUs) and new parallel programming models, namely Partitioned Global Address Space (PGAS) and message passing. The message passing paradigm, particularly the Message Passing Interface (MPI), is the prevailing method for parallel programming today, however, Partitioned Global Address Space (PGAS) is the close competitor. The ease of use in PGAS due to the nice abstract view comes at a price that makes MPI remaining

to be a solid competitor. However, Cray is one of the few vendors who offer on their platforms two leading PGAS languages, UPC and Chapel. With the new supercomputer, we are planning to conduct extensive productivity comparative studies aiming at Chapel, UPC, and MPI. Productivity will be assessed not simply based on the number of lines of codes and execution time. Instead, in PGAS for example, our work will target and address translation overheads associated with the PGAS memory model, lack or efficient compiler optimizations due to suspected pointer aliasing, and synchronization to name a few. Even ease of use, will consider establishing metrics and understanding of conceptual ease of use problems beyond the number of lines of codes, including the ability to express and deal with large data problems including some of the random memory accesses types of problems. Workloads will be selected from benchmarking suites like the NAS Parallel Benchmark and the High-Performance Computing Challenge Benchmark (HPCC).

The Cray compilers come with low level software and hardware optimizations for such programming languages and therefore, they form perfect testbeds for parallel programming studies.

The Co-PI research area is related to database and data processing, which focuses on voluminous data sets (e.g. the 80TB Common Crawl Corpus, the 2.2TB Google Books Ngrams, the 2.2TB Google Books Ngrams, from Stanford). Since the datasets cannot be processed using any single computer, the purpose of this research is to investigate an infrastructure of using a parallel computing system for pervasive multi-dimensional spatial data sharing and access. Despite the fact that considerable research has been done on conventional data access, there has been little work done in integrating content-based multi-dimensional data in the pervasive computing environment, especially in wireless networks. In addition, there is not much research work reported on the semantic analysis and content representation of multi-dimensional data. These research issues, however, are crucial for successful and efficient information system applications such as GIS, gene expression analysis, social network modeling, and multimedia information retrieval. Therefore, it is highly necessary to investigate these challenges and devise a novel methodology for multi-dimensional data integration.

The second Co-PI is conducting research on DNA barcoding of tropical species in collaboration with Godfrey Okoye University, Enugu, Nigeria with technical support from the DNA Learning Center, Cold Spring Harbor Laboratory, NY, USA. The barcode sequence data generated from plants, animals, fungi and some bacteria from Eastern Nigeria will be analyzed for sequence similarities to determine species identity, diversity and distribution in the ecosystem. The outcome of this project will help pharmaceutical industries, plant and animal breeders, nature conservationists and all other users of natural resources to properly identify and use the biological organisms that are native to Eastern Nigeria. This effort will also lead to the discovery and proper cataloging of new species that till this moment have not been documented or studied.

The database for the sequenced data is kept at Bowie State University where the sequences will be analyzed using a high speed computer for nucleotide sequence differences and alignment between the species from Eastern Nigeria and the other sequences in genebanks around the world.

It is expected that this research will lead to an accumulation of very large volumes of DNA sequence data that will take very long to align and compare with already existing sequences in several DNA genebanks around the world if we used a regular speed computer. It will take months to analyze the data which otherwise would have taken hours to do if we have a high power or a super computer. The supercomputer will be used for faster inferences and completion of the project.

In addition, Bowie State University (BSU) is a historically black university (HBCU) which educates and trains the next generation of black leaders. Using the supercomputer, we will be assigning realistic computational problems that can integrate research and teaching in the STEM disciplines. The supercomputer will be integrated in a number of our course offerings. A plan is also being devised to have more workshops and summer training to the faculty and to the students to stimulate research in engineering, science and mathematics through computational modeling and simulations. Work will be widely disseminated through standard academic venues.